

Sensor Data Abstraction Using Flexible Layered Slope Grid

Young Jin Jung¹, Seon-Phil Jeong^{2*}, Moon Sun Shin³ and Hyunmi Choi⁴

¹Agency for Defense Development, Korea

²Division of Technology and Science, BNU-HKBU United International College,
China

³Dept. of Computer Engineering, Konkuk University, Korea

⁴Dept. of Advanced Application Environment, Development, KISTI, Koear

yjjung@add.re.kr, spjeong@uic.edu.hk, msshin@kku.ac.kr, hmchoi@kisti.re.kr

*Corresponding Author

Abstract

An environmental monitoring application is designed to monitor and track various kinds of environmental phenomena such as water pollution and global warming. By collecting detected sensor data from the sensors installed at a target place, a monitoring system analyzes and predicts an environmental change. For smoothly answering a user query, a data abstraction model is designed to rapidly process an amount of sensor data by employing data filtering, data aggregation, and data summarization. In order to dynamically represent the different layer of sensor data in a region, we design FLSA (Flexible Layered Sensor Data Abstraction), which supports a flexible layer of SGSA (Slope Grid for Sensor Data Abstraction). When the sensor data is transmitted to a server, FLSA make some clusters to shows several layer of sensor data depending on sensor data gradient. Each cell of FLSA has a set of SGSA to represent each sensor data layer in the cell. This layered SGSA of FLSA is useful to represent a distributed sensor data in a cell. When an abstraction model is used to environmental monitoring, there is a trade-off between faster data processing and data representation. FLSA focuses on data representation to show detailed condition of a target place.

Keywords: Sensor data abstraction, Layered abstraction, Flexible layered abstraction, Slope grid, SGSA

1. Introduction

In order to understand a change of condition of a remote place, an environmental monitoring application is designed to collect and analyze various kinds of sensor data. The received sensor data is summarized and aggregated to make a context to show a useful data or information. When a user issues a query to get a particular data, a system processes an amount of sensor data to answer the query. However, some queries make a system demand more time to process data and to derive a high context from raw sensor data. There are some examples such as “where is an area to show rapidly changed water pollution level?” and “Track water pollution areas and their pollution level for a year”.

To reduce a cost to answer a user query, we design FLSA (Flexible Layered Sensor Data Abstraction) to support dynamic sensor data layers. To represent each sensor data layer, FLSA employs SGSA (Slope Grid for Sensor Data Abstraction) [1], which has a slope grid to summarize the data. This abstraction model represents a data trend of a cell on a grid. To show a data trend of each cell, SGSA uses a slope grid to simply represent sensor data in a cell by defining a slope direction. When a system processes a query, it checks the direction and the boundary of SGSA to extract a particular area, that has a requested value. After collecting the set of SGSA for a particular area, a system makes a context to represent overall data trend of the extracted SGSA. A system can reduce the cost to make an answer for a user query by examining the data trend of SGSA instead of accessing the raw sensor data.

To extend the representation of SGSA, LSGSA is designed to represent the separated layered slopes, however there is a limit of layers. When a cell of a grid has more sensors, SGSA and LSGSA is not enough to represent a dynamic data trend, which has different data level. FLSA can represent more sensor data with layered SGSA to summarize each data cluster.

2. Related Work

There are various kinds of environmental monitoring applications such as a habitat monitoring application [2], GoMOOS (Gulf of Maine Ocean Observing System) [3], PODS, SIFTS (Short-term Inundation Forecasting for Tsunamis) [4], PORTS (Physical Oceanographic Real-Time System) [5], CMOP (Center for Coastal Margin Observation & Prediction) [6], FloodNet [7], and an air pollution monitoring [8]. These monitoring applications deal with sensor data to understand the change of condition with domain knowledge.

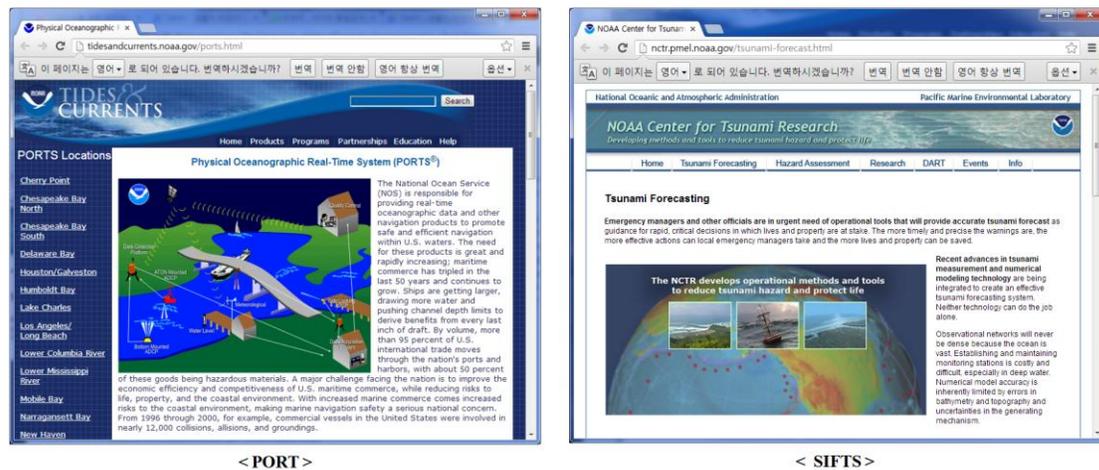


Figure 1. Environmental monitoring applications

Great Duck Island (GDI) system monitors a particular bird such as stormy petrel by installing sensors on an island [2]. GoMOOS collects and analyzes sensor data from

buoys and stations to understand the change of ocean condition such as currents, waves, wind, and HF radar [3].

SIFTS is designed for emergency managers and officials to provide a fast tsunami forecast service. To support an effective tsunami forecasting service, tsunami measurement and numerical modeling technology are integrated [4]. CMOP shows a new paradigm for conducting coastal-margin science by a structured integration of information [6]. FloodNet is designed to support the functional floodplain conditions at a particular location [7]. An air pollution monitoring shows the air pollution (dust) prevention scenario for a building manger by using an air pollution forecasting with sensor data [8].

Data abstraction is used to effectively process queries or to derive a high level context [1, 9, 10, 11, 12]. There are effective abstraction methods such as MauveDB [13, 14], SAIL (Sensor Abstraction and Integration Layers) [11, 15], OSGi service platform [18], and an abstraction of regions [10], SGSA [1], LSGSA [9].

MauveDB shows an abstraction as a model based views. An additional layer for abstraction between received sensor data and a higher view is used to represent an observation by using an approximation model. The layered architecture of SAIL is used to collect observation data and to support node interaction between WSN applications and context-aware systems [11]. This architecture includes access, abstraction, and integration layers to support sensor network discovery and event monitoring in OSGi service platform. An abstraction of regions is designed to derive higher context by encapsulating low-level data communication, resource usage, and data sharing on local regions.

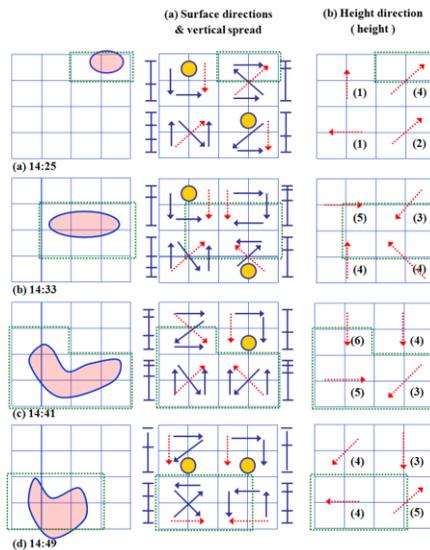


Figure 2. Abstraction data update (LSGSA)

SGSA is designed for rapidly processing data by using a simple slope grid on a local area. LSGSA is an extended version of SGSA to improve the presentation of SGLA by employing layered slopes such as a surface slope to show the maximum data trend and a

height slope to represent an inner condition on a cell of a grid. These effective environmental monitoring applications and abstraction models on sensor network are effective to derive higher context or useful data from raw sensor data.

3. Flexible Layered Sensor Data Abstraction

In an environmental monitoring application, sensor data abstraction is useful to handle an amount of data. When a user issues a query to ask about the condition of a target place, a monitoring application makes useful information from the detected sensor data. To rapidly answer a user query, our abstraction model is used to quickly check sensor data to satisfy a user's query by summarizing sensor data in a cell.

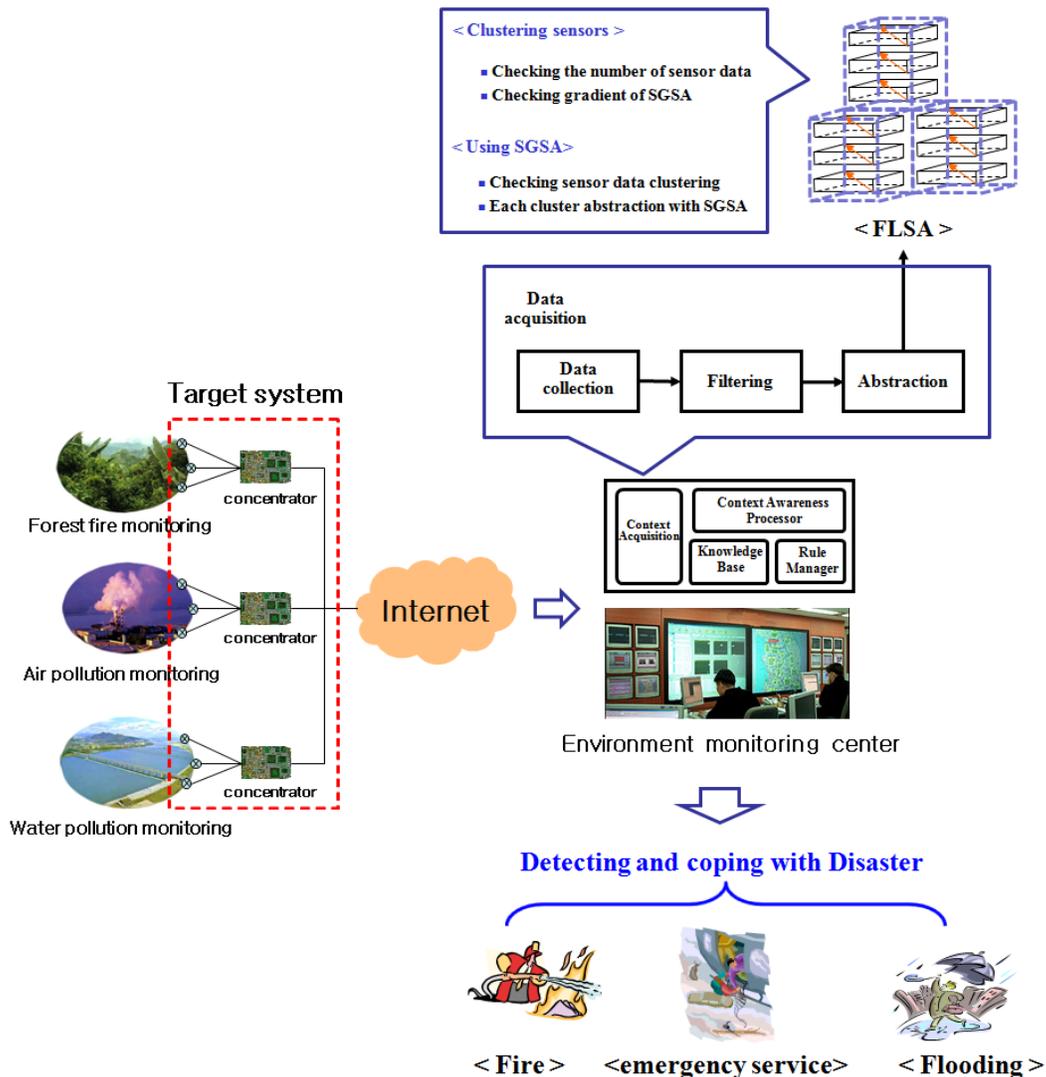
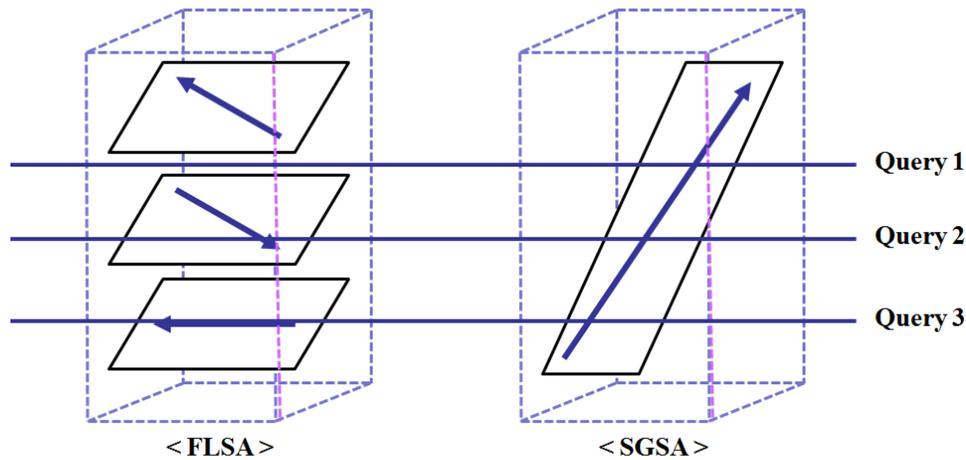


Figure 3. Sensor data abstraction with FLSA

Figure 3 shows the data processing of data acquisition module to make FLSA in the environmental monitoring center. Observation data is transmitted to a monitoring server from the installed sensors on the target place such as a forest, a factory, a lake, a river, etc. On data acquisition module, sensor data is collected, filtered, and abstracted. To represent raw sensor data for each cell of a grid, we use FLSA, which is useful to show layered sensor data.



< Selected data to answer queries after checking SGSA >

	FLSA	SGSA
Query 1	null	All of sensor data
Query 2	Cluster 2	All of sensor data
Query 3	Cluster 3	All of sensor data

Figure 4. Comparison between FLSA and SGSA

Figure 4 show the difference between FLSA and SGSA. SGSA is used to abstract raw sensor data with a simple slope grid. When sensor data, which has significantly different value, is abstracted, a gradient of SGSA is heavily tilted, because the slope grid has to include all of the raw sensor data on a cell. To answer queries of Figure 4, the cell of SGSA is selected and all of the included sensor data is examined. In FLSA case, the cell of FLSA is also selected to answer the queries of Figure 4. FLSA can reduce data access cost, because FLSA can check the boundary of the included SGSA instead of raw sensor data. In the example of queries of figure 4, FLSA returns the selected data for queries after checking the boundary of the included SGSA such as null for query 1, cluster 2 for query 2, and cluster 3 for query 3.

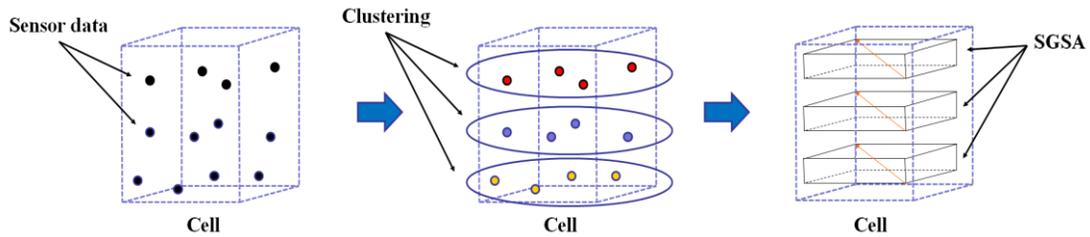


Figure 5. Sensor data abstraction with FLSA

In FLSA, sensor data of a cell, which means a local area, is classified depending on the gradient of each cluster as shown in Figure 4 and Figure 5. With this clustering, each classified sensor data layer has a minimized boundary, which is useful to reduce the data access cost to answer a query.

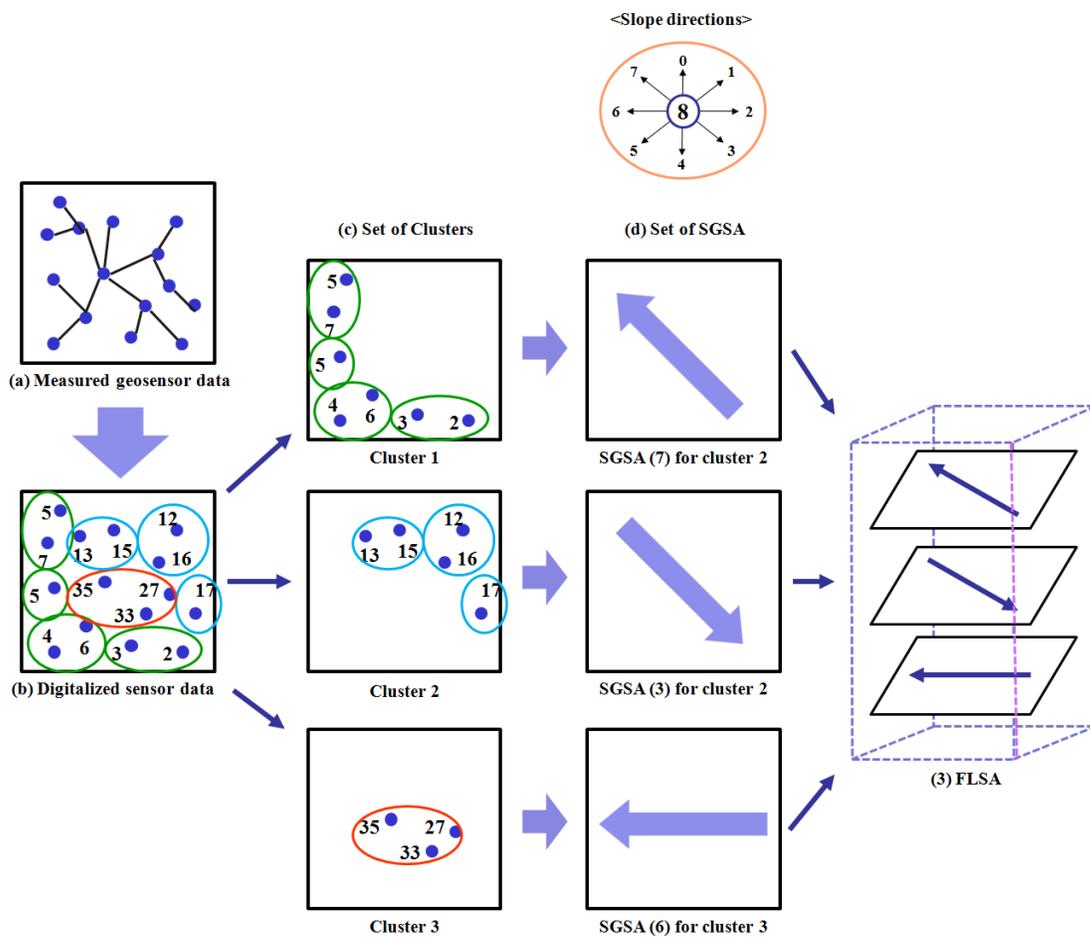


Figure 6. Sensor data abstraction with FLSA

Figure 6 shows data abstraction step of FLSA from observation data to FLSA. To make a cluster, sensor data is classified depending on the decreased gradient of each cluster as shown in algorithm 1. A slope direction is derived for each cluster at (d) set of SGSA according to <slope direction> [1].

Algorithm 1 shows sensor data clustering and abstraction of FLSA. Observed sensor data is sorted as a first step of clustering. From the maximum sensor data, FLSA makes a cluster. When data is inserted into a cluster, FLSA check the gradient change of the cluster with a limit of cluster such as 10° or 15°. The limit of cluster is generally defined by a manager of environmental monitoring application. This limitation can be changed depending on situation of target place such as an emergency. As a result, each cluster has a minimized boundary within a limitation. The slope direction for each cluster is also derived by using SGSA algorithm [1].

Algorithm 1. Clustering of FLSA

```

Algorithm CellOfFLSA::ClusteringOnCell (Cell *cell , SensorData *data[] )
input : cell           // a cell of FLSA
output : cell          // the cell, which has layers of SGSA depending on clustering
method :
    // clustering of sensor data on a cell

    FLSA_cluster clusterOfFLSA;           // FLSA cluster for each cell
    Sensor_data[]= sort(data[]);          // sensor data is sorted
    temp_cluster = new cluster();

    While (!sensor_data.isEmpty()){

        temp_gradient = temp_cluster.checkGradient();    // get the gradient of cluster

        // put maximum sensor data to cluster from SensorData
        temp_cluster.insertData(sensor_data.getMaxData);

        gradient = temp_cluster.checkGradient();    // get the gradient of cluster

        if(limitOfCluster(gradient, temp_gradient)){ // limit of cluster to check change of gradient
            temp_cluster.removeLastData();
            temp_cluster.GetDirectionOfSGSA ();    // define a slope direction of a cluster

            clusterOfFLSA.put(temp_cluster);      // add temp_cluster to FLSA

            Temp_cluster = new cluster();         // new cluster
        }
        else{ // when sensor data inserted, the gradient of the cluster is satisfied
            sensor_data.removeMaxData();
        }
    }

    Cell.cluster = clusterOfFLSA;
    Cell.defineBoundary();                       // define the boundary to include FLSA clusters

return cell // the cell, which has layered SGSA
end

```

4. FLSA Utilization

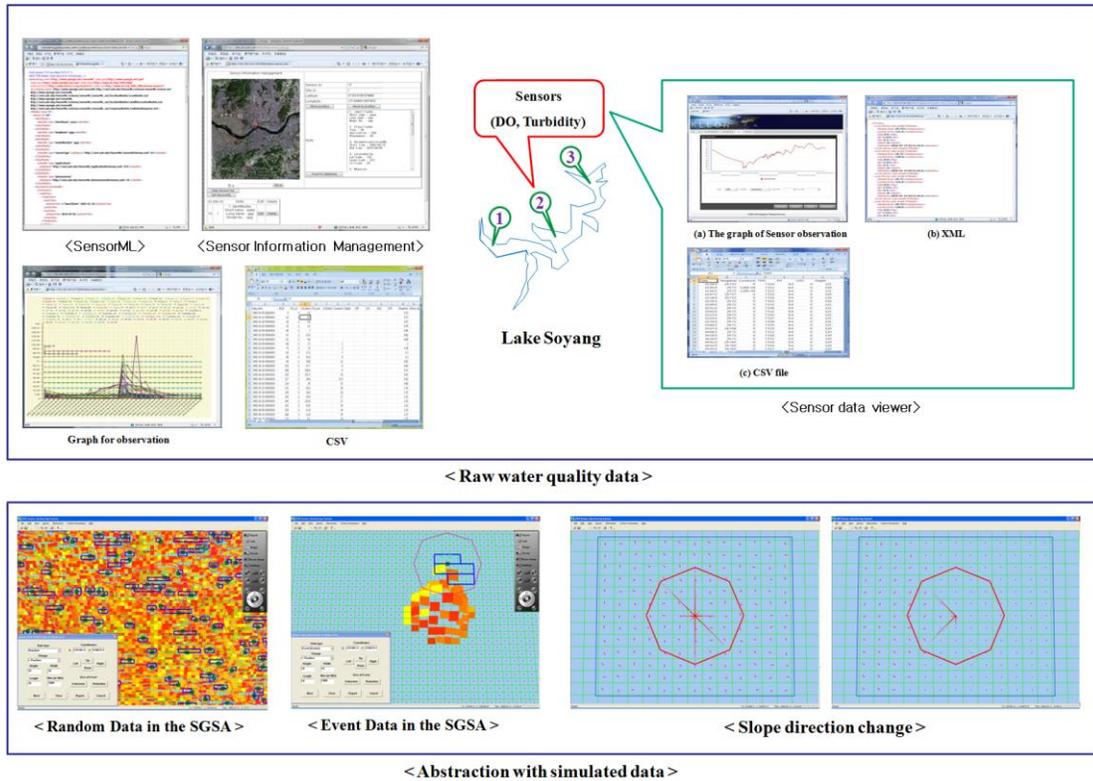


Figure 7. Observation data management and abstraction utilization for water quality monitoring application

To utilize FLSA on water quality monitoring application, we are implementing an observation data processing module, which includes an observation data management (graph, file), an observation site management, and FLSA on water quality monitoring application. Water quality data is distributed spatially, temporally, and vertically depending on observation depth of the lake or the river. To deal with sensor data, we implemented observation site management page, which employs sensor information management with SensorML [17]. Users search and check observation data with graphs (a temporal graph and a depth graph) and the CSV file. To test data abstraction with SGSA, we use simulated data instead of raw sensor data. We can extract higher value from raw simulated data to answer a query. To show an overall data trend at a particular area, a slope direction summarization is also derived from SGSA. We are implementing FLSA, which is a set of SGSA, to support layered sensor data abstraction. To support a user friendly interface, these data processing and data abstraction methods can be utilized on web based simulation service [18].

7. Conclusion

In this paper, we design FLSA to support dynamic sensor data clustering and abstraction depending on the difference between sensor data on a cell, which means a local area. When

we supports layered sensor data abstraction with FLSA, data access cost to answer a query will decrease as shown in Figure 4. FLSA employs SGSA to represent each data layer that is a cluster. The boundary of each cluster can be changed to keep a minimized boundary, which is suitable for a situation of a target place, depending on the limitation of the cluster. This limitation is generally defined by a manager of environmental monitoring application. This dynamic data abstraction layer will be used to support faster data processing on environmental monitoring applications.

References

- [1] Y. J. Jung and S. Nittel, "Geosensor Data Abstraction for Environmental Monitoring Application", Proceedings of the 5th international conference on Geographic Information Science, Published in: Lecture Notes in Computer Science, Springer, Berlin, vol. 5266, (2008), pp. 168-180, doi: 10.1007/978-3-540-87473-7_11.
- [2] D. Culler, J. Polastre, R. Szewczyk and J. Anderson, "Wireless sensor networks for habitat monitoring", WSN, (2002), pp. 88-97.
- [3] Gulf of Maine Ocean Observing System (GoMOOS), <http://tidesandcurrents.noaa.gov/gomoos.html>.
- [4] Short-term Inundation Forecasting for Tsunamis (SIFT) system, (2006) May, NOAA center for tsunami, <http://nctr.pmel.noaa.gov/tsunami-forecast.html>.
- [5] Physical Oceanographic Real-Time System (PORTS), the Center for Operational Oceanographic Products and Services (CO-OPS), (2009) October, <http://co-ops.nos.noaa.gov/ports.html>.
- [6] Project I.4.3: Model services and interfaces, Center for Coastal Margin, Observation & Prediction (CMOP), <http://www.ccalmr.ogi.edu/research/collaboratories/projects/modelingservices>, (2009).
- [7] J. Zhou and D. De Roure, "FloodNet: Coupling Adaptive Sampling with Energy Aware, Routing in a Flood Warning System", Journal of Computer Science and Technology, vol. 22, no. 1, (2007), pp. 121-130.
- [8] Y. J. Jung, Y. K. Lee, D. G. Lee, Y. M. Lee, S. Nittel, K. Beard, K. W. Nam and K. H. Ryu, "Design of Sensor Data Processing Steps in an Air Pollution Monitoring System", Sensors, vol. 11, no. 12, (2011), pp. 11235-11250; doi:10.3390/s111211235.
- [9] Y. M. Lee, Y. J. Jung, K. W. Nam, S. Nittel, K. Beard and K. H. Ryu, "Geosensor Data Representation Using Layered Slope Grids," Sensors 2012, vol. 12, no. 12, pp. 17074-17093 (2012); doi:10.3390/s121217074.
- [10] M. Welsh, G. Mainland, "Programming Sensor Networks Using Abstract Regions," Proceedings of the 1st conference on Symposium on Networked Systems Design and Implementation, vol. 1, USENIX Association, Berkeley, CA, USA, (2004), pp. 3-3.
- [11] M. Girolami, S. Lenzi, F. Furfari and S. Chessa, "SAIL: a Sensor Abstraction and Integration Layer for Context Awareness", Proceedings of the 2008 34th Euromicro Conference Software Engineering and Advanced Applications, (2008), pp. 374-381, doi: 10.1109/SEAA.2008.30.
- [12] K. Aberer, M. Hauswirth and A. Salehi, "A Middleware For Fast And Flexible Sensor Network Deployment", 32nd Very Large Data Base (VLDB), (2006), pp. 1199-1202.
- [13] A. Deshpande and S. Madden, "MauveDB: Supporting Model based User Views in Database Systems", Proceedings of the 2006 ACM SIGMOD international conference on Management of data, Chicago, IL, USA, (2006), pp. 73-84, doi: <http://doi.acm.org/10.1145/1142473.1142483>.
- [14] B. Kanagal and A. Deshpande, "Online Filtering, Smoothing and Probabilistic Modeling of Streaming data", Proceedings of the 2008 IEEE 24th International Conference on Data Engineering, (2008), pp. 1160-1169, doi: 10.1109/ICDE.2008.4497525
- [15] G. Gigan and I. Atkinson, "Sensor Abstraction Layer: a unique software interface to effectively manage sensor networks", Proceedings of the 2007 International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Melbourne, Australia, (2007), pp. 479-484, doi: 10.1109/ISSNIP.2007.4496890.
- [16] OSGi Service Platform Release 4, Version 4.1, (2007) May, Open Standards Organization (OSGi), <http://www.osgi.org/Release4/HomePage>.
- [17] M. Botts, "Sensor Web Enablement", <http://www.opengeospatial.org/>.
- [18] Y. J. Jung, J. B. Moon, D. S. Jin, B. Y. Ahn, J. H. Seo, H. Ryu, O. H. Byeon and J. S. R. Lee, "Performance Improvement for Web based Simulation Service on EDISON_CFD", International Journal of Software Engineering and Its Applications, to be accepted.

Authors



Dr. Youngjin Jung

He is a researcher at Agency for Defense Development of Korea. He got his Ph.D. degree in computer science from Chungbuk National University, Korea. His research interests include data provenance, spatiotemporal database application, moving object index, and environmental monitoring application



Dr. Seonphil Jeong

Currently, he is an assistant professor at BNU-HKBU United International College in Zhuhai, China. He received his Ph.D. degree in MIS from Chungbuk National University, Korea. His research interests are system analysis and design, Project Management, IS auditing and MIS.



Dr. Moon Sun SHIN

She received her Ph. D. degree at the Department of Computer Engineering from Chungbuk National University in 2004. She is currently working as a professor in Konkuk University since 2005. Her research interests include sensor data processing and database security



Ms. Hyunmi Choi

She is a researcher at the Department of Overseas Information, Korea Institute of Science and Technology Information, Daejeon. She received her BA in Computer Engineering at Hanbat National University in 2002, and her MS in Computer Science from the Chungbuk National University in 2004, respectively. Her research interests include Moving Object System, Information Retrieval System, and Database System.